1979

Simulation of Nonhomogeneous Poisson Processes by Thinning

Lewis, P.A.W.; Shedler, G.S.

Lewis, PA W., and Gerald S. Shedler. "Simulation of nonhomogeneous Poisson processes by thinning." Naval research logistics quarterly 26.3 (1979): 403-413.
http://hdl.handle.net/10945/63159

SIMULATION OF NONHOMOGENEOUS POISSON PROCESSES BY THINNING

P. A. W. Lewis*

Naval Postgraduate School
Monterey, California

G. S. Shedler

IBM Research Laboratory
San Jose, California

ABSTRACT

A simple and relatively efficient method for simulating one-dimensional and two-dimensional nonhomogeneous Poisson processes is presented. The method is applicable for any rate function and is based on controlled deletion of points in a Poisson process whose rate function dominates the given rate function. In its simplest implementation, the method obviates the need for numerical integration of the rate function, for ordering of points, and for generation of Poisson variates.

1. INTRODUCTION

The one-dimensional nonhomogeneous Poisson process (see e.g. [6], pp. 28-29; [4], pp. 94-101) has the characteristic properties that the numbers of points in any finite set of nonoverlapping intervals are mutually independent random variables, and that the number of points in any interval has a Poisson distribution. The most general nonhomogeneous Poisson process can be defined in terms of a monotone nondecreasing right-continuous function $\Lambda(x)$ which is bounded in any finite interval. Then the number of points in any finite interval, for example $(0, x_0]$, has a Poisson distribution with parameter $\mu_0 = \Lambda(x_0) - \Lambda(0)$. In this paper it is assumed that $\Lambda(x)$ is continuous, but not necessarily absolutely continuous. The right derivative $\lambda(x)$ of $\Lambda(x)$ is called the rate function of the process; $\Lambda(x)$ is called the integrated rate function and has the interpretation that for $x \geqslant 0$, $\Lambda(x) - \Lambda(0) = E[N(x)]$, where $N(x)$ is the total number of points in $(0, x]$. Note that $\lambda(x)$ may jump at points at which $\Lambda(x)$ is not absolutely continuous. In contrast to the homogeneous Poisson process, i.e., $\lambda(x)$ is a constant (usually denoted by λ), the intervals between the points in a one-dimensional nonhomogeneous Poisson process are neither independent nor identically distributed.

Applications of the one-dimensional nonhomogeneous Poisson process include modelling of the incidence of coal-mining disasters [6], the arrivals at an intensive care unit [12], transaction processing in a data base management system [14], occurrences of major freezes in Lake

*Support under Office of Naval Research Grant NR-42-343 is gratefully acknowledged.

Constance [20], and geomagnetic reversal data [19]. The statistical analysis of trends in a one-dimensional nonhomogeneous Poisson process, based on the assumption of an exponential polynomial rate function, is discussed by [6], [5], [12], and [14].

There are a number of methods for simulating the nonhomogeneous Poisson process which we review briefly.

(i) Time-scale transformation of a homogeneous (rate one) Poisson process via the inverse of the (continuous) integrated rate function $\Lambda(x)$ constitutes a general method for generation of the nonhomogeneous Poisson process (cf., [4], pp. 96-97). This method is based on the result that X_1, X_2, ..., are the points in a nonhomogeneous Poisson process with continuous integrated rate function $\Lambda(x)$ if and only if $X_1' = \Lambda(X_1)$, $X_2' = \Lambda(X_2)$, ..., are the points in a homogeneous Poisson process of rate one. The time-scale transformation method is a direct analogue of the inverse probability integral transformation method for generating (continuous) nonuniform random numbers. For many rate functions, inversion of $\Lambda(x)$ is not simple and must be done numerically; cf., [7] and [17]. The resulting algorithm for generation of the nonhomogeneous Poisson process may be far less efficient than generation based on other methods; see e.g., [13], [15], and [17] for discussions of special methods for efficiently generating the nonhomogeneous Poisson process with log-linear and log-quadratic rate functions.

(ii) A second general method for generating a nonhomogeneous Poisson process with integrated rate function $\Lambda(x)$ is to generate the intervals between points individually, an approach which may seem more natural in the event-scheduling approach to simulation. Thus, given the points $X_1 = x_1$, $X_2 = x_2$, ..., $X_i = x_i$, with $X_1 < X_2 < \ldots < X_i$, the interval to the next point, $X_{i+1} - X_i$, is independent of $x_1, \ldots, x_{i-1}$ and has distribution function $F(x) = 1 - \exp[-\{\Lambda(x_i + x) - \Lambda(x_i)\}]$. It is possible to find the inverse distribution function $F^{-1}(\cdot)$, usually numerically, and generate $X_{i+1} - X_i$ according to $X_{i+1} - X_i = F^{-1}(U_i)$, where U_i is a uniform random number on the interval (0,1]. Note, however, that this not only involves computing the inverse distribution function for each interval $X_{i+1} - X_i$, but that each distribution has different parameters and possibly a different form. An additional complication is that $X_{i+1} - X_i$ is not necessarily a proper random variable, i.e., there may be positive probability that $X_{i+1} - X_i$ is infinite. It is necessary to take this into account for each interval $X_{i+1} - X_i$ before the inverse probability integral transformation is applied. The method is therefore very inefficient with respect to speed, more so than the time-scale transformation method.

(iii) In a third method, simulation of a nonhomogeneous Poisson process in a fixed interval $(0, x_0]$ can be reduced to the generation of a Poisson number of order statistics from a fixed density function by the following result (cf., [6], p. 45). If $X_1, X_2, \ldots, X_n$ are the points of the nonhomogeneous Poisson process in $(0,x_0]$, and if $N(x_0) = n$, then conditional on having observed $n(> 0)$ points in $(0, x_0]$, the X_i are distributed as the order statistics from a sample of size n from the distribution function $\{\Lambda(x) - \Lambda(0)\}/\{\Lambda(x_0) - \Lambda(0)\}$, defined for $0 < x \leqslant x_0$. Generation of the nonhomogeneous Poisson process based on order statistics is in general more efficient (with respect to speed) than either of the previous two methods. Of course, a price is paid for this greater efficiency. First, it is necessary to be able to generate Poisson variates, and second, more memory is needed than in the interval-by-interval method in order to store the sequence of points. Enough memory must be provided so that with very high probability the random numbers of points generated in the interval can be stored. Recall that the number of points in the interval $(0, x_0]$ has a Poisson distribution with mean

$\mu_0 = \Lambda(x_0) - \Lambda(0)$. Memory of size, e.g., $\mu_0 + 4\mu_0^{1/2}$ wll ensure that overflow will occur on the average in only one out of approximately every 40,000 realizations. This probability is small enough so that in the case of overflow, the realization of the process can generally be discarded.

(iv) Again, there is a very particular and very efficient method for simulation of nonhomogeneous Poisson processes with log-linear rate function [13] which, at the cost of programming complexity and memory, can be used to obtain an efficient simulation method for other rate functions, as in [15].

In this paper a new method is given for simulating a nonhomogeneous Poisson process which is not only conceptually simple, but is also computationally simple and relatively efficient. In fact, at the cost of some efficiency, the method can be applied to simulate the given nonhomogeneous Poisson process *without the need for numerical integration or routines for generating Poisson variates.* Used in conjunction with the special methods given in [13] and [15], the method can be used to generate quite efficiently nonhomogeneous Poisson processes with rather complex rate functions, in particular combinations of long-term trends and fixed-cycle effects. The method is also easily extended to the problem of generating the two-dimensional nonhomogeneous Poisson process.

2. SIMULATION OF ONE-DIMENSIONAL NONHOMOGENEOUS POISSON PROCESSES

Simulation of a nonhomogeneous Poisson process with general rate function $\lambda(x)$ in a fixed interval can be based on thinning of a nonhomogeneous Poisson process with rate function $\lambda^*(x) \geqslant \lambda(x)$. The basic result is

THEOREM 1: Consider a one-dimensional nonhomogeneous Poisson process $\{N^*(x):x \geqslant 0\}$ with rate function $\lambda^*(x)$, so that the number of points, $N^*(x_0)$, in a fixed interval $(0, x_0]$ has a Poisson distribution with parameter $\mu_0^* = \Lambda^*(x_0) - \Lambda^*(0)$. Let $X_1^*, X_2^*, \ldots, X_{N^*(x_0)}^*$ be the points of the process in the interval $(0,x_0]$. Suppose that for $0 \leqslant x \leqslant x_0$, $\lambda(x) \leqslant \lambda^*(x)$. For $i = 1, 2, \ldots, n$, delete the point X_i^* with probability $1 - \lambda(X_i^*)/\lambda^*(X_i^*)$; then the remaining points form a nonhomogeneous Poisson process $\{N(x):x \geqslant 0\}$ with rate function $\lambda(x)$ in the interval $(0,x_0]$.

PROOF: Since $\{N^*(x):x \geqslant 0\}$ is a nonhomogeneous Poisson process and points are deleted independently, it is clear that the number of points in $\{N(x):x \geqslant 0\}$ in any set of nonoverlapping intervals are mutally independent random variables. Thus, it is sufficient to show that the number of points $N(a,b)$ in $\{N(x):x \geqslant 0\}$ in an arbitrary interval $(a,b]$ with $0 \leqslant a < b \leqslant x_0$ has a Poisson distribution with parameter $\Lambda(b) - \Lambda(a)$. Observe that with $p(a,b) = \{\Lambda(b) - \Lambda(a)\}/\{\Lambda^*(b) - \Lambda^*(a)\}$, we have the conditional probability:

$$(1) \qquad P\{N(a,b) = n \mid N^*(a,b) = k\} = \begin{cases} 1 & \text{if } n = k = 0 \\ \binom{k}{n} \{p(a,b)\}^n \{1 - p(a,b)\}^{k-n} & \text{if } k \geqslant n \geqslant 0 \\ & \text{and } k \geqslant 1 \\ 0 & \text{if } n \geqslant 1 \\ & \text{and } k < n \end{cases}$$

Equation (1) is a consequence of the well-known result that, conditional on k (>0) points in the interval $(a,b]$, the joint density of the k points in the process $\{N^*(x):x \geqslant 0\}$ is $\lambda^*(x_1) \ldots \lambda^*(x_k)/\{\Lambda^*(b) - \Lambda^*(a)\}^k$. The desired result is obtained in a straightforward manner from Equation (1) by removing the condition.

Theorem 1 is the basis for the method of simulating nonhomogeneous Poisson processes given in this paper.

ALGORITHM 1: One-dimensional nonhomogeneous Poisson process.

1. Generate points in the nonhomogeneous Poisson process $\{N^*:(x) \geqslant 0\}$ with rate function $\lambda^*(x)$ in the fixed interval $(0,x_0]$. If the number of points generated, n^*, is such that $n^* = 0$, exit; there are no points in the process $\{N(x):x \geqslant 0\}$.

2. Denote the (ordered) points by $X_1^*, X_2^*, \ldots, X_{n^*}^*$. Set $i = 1$ and $k = 0$.

3. Generate U_i, uniformly distributed between 0 and 1. If $U_i \leqslant \lambda(X_i^*)/\lambda^*(X_i^*)$, set k equal to $k+1$ and $X_k = X_i^*$.

4. Set i equal to $i+1$. If $i \leqslant n^*$, go to 3.

5. Return $X_1, X_2, \ldots, X_n$, where $n = k$, and also n.

(i) In the case where $\{N^*(x):x \geqslant 0\}$ is a homogeneous Poisson process with $\lambda^*(x) = \lambda^*$;

(ii) the minimum of $\lambda(x)$, say $\underline{\lambda}$, is known, and

(iii) generation of uniformly distributed variates is computationally costly,

considerable speedup can be obtained by noting that X_i^* is always accepted if $U_i \leqslant \underline{\lambda}/\lambda^*$. This obviates, in some cases, computation of $\lambda(x)$, which is the main source of inefficiency in the algorithm. Moreover, in this case $\lambda^* U_i/\underline{\lambda}$ can be used as the next uniformly distributed variate.

The method of thinning in this simple form, i.e., $\lambda^*(x) = \lambda^* = \max_{0 \leqslant x \leqslant x_0} \lambda(x)$, can also be used to provide an algorithm for generating a nonhomogeneous Poisson process on an interval-by-interval basis, as discussed in subsection (ii) of Section 1. The interval to the next point $X_{i+1} - X_i$ is obtained by generating and cumulating exponential (λ^*) random numbers $E_1^*, E_2^*, \ldots,$ until for the first time $U_j \leqslant \lambda(X_i + E_1^* + \ldots E_j^*)/\lambda^*$, where the U_j are independent uniform random numbers between 0 and 1. This algorithm is considerably simpler than the interval-by-interval algorithm of Section 1 since it requires no numerical integration, only the availability of uniform random numbers.

3. DISCUSSION OF THE METHOD OF THINNING

(i) Relationship to acceptance-rejection method

The method of thinning of Algorithm 1 is essentially the obverse of the conditional method of Section 1, using conditioning and acceptance-rejection techniques to generate the random variables with density function $\lambda(x)/\{\Lambda(x) - \Lambda(0)\}$ (Lewis and Shedler, [15], Algorithm 3). The differences are subtle, but computationally important. In the acceptance-rejection method, it is first necessary to generate a Poisson variate with mean $\mu_0 = \Lambda(x_0) - \Lambda(0)$, and this involves an integration of the rate function $\lambda(x)$. Then the Poisson (μ_0) number, n, of variates generated by acceptance-rejection must be ordered to give $X_1, X_2, \ldots, X_n$.

(ii) Simplest form of the thinning algorithm

In the simplest form of the method of thinning, $\lambda^*(x)$ is taken to be a constant λ^*, so that, for instance, the points $X_1^*, X_2^*, \ldots, X_{n^*}^*$ can be generated by cumulating exponential (λ^*) variates until the sum is greater than x_0 (cf., [13], Algorithm 1). Thinning is then applied to the generated points. *No ordering, no integration of* $\lambda(x)$ *and no generator of Poisson variates is required.* Of course for both algorithms to be efficient, computation of $\lambda(x)$ and $\lambda^*(x)$ must be easy relative to computation of the inverse of $\Lambda(x)$.

(iii) Efficiency

For the thinning algorithm (as well as the algorithm based on conditioning and acceptance-rejection) efficiency, as measured by the number of points deleted, is proportional to $\mu_0/\mu_0^* = \{\Lambda(x_0) - \Lambda(0)\}/\{\Lambda^*(x_0) - \Lambda^*(0)\}$; this is the ratio of the areas between 0 and x_0 under $\lambda(x)$ and $\lambda^*(x)$. Thus, $\lambda^*(x)$ should be as close as possible to $\lambda(x)$ consistent with ease of generating the nonhomogeneous Poisson process $\{N^*(x):x \geqslant 0\}$.

(iv) An example: fixed cycle plus trend

To illustrate the applicability of the thinning algorithm, consider its use in conjunction with the algorithms given by [13] and [15] for log-linear and log-quadratic rate functions. Assume that it is necessary to simulate a nonhomogeneous Poisson process whose rate function increases quadratically with time but also has a fixed-period cycle, e.g.,

$$\lambda(x) = \exp\{\alpha_0 + \alpha_1 x + \alpha_2 x^2 + K \sin(\omega_0 x + \theta)\},$$
$$0 \leqslant x \leqslant x_0;\ K \geqslant 0;\ 0 < \theta \leqslant 2\pi;\ \omega_0 > 0.$$

This is the model found by Lewis [12] for arrivals at an intensive care unit, where there is a strong time-of-day effect. Thus if $\omega_0 = 2\pi/T_0$, then the period $T_0 = 1$ day. Computation of $\Lambda^{-1}(\cdot)$ is difficult. To determine $\lambda^*(x)$, note that

$$\lambda(x) \leqslant \lambda^*(x) = \exp\{\alpha_0 + K + \alpha_1 x + \alpha_2 x^2\},$$

and therefore

$$\lambda(x)/\lambda^*(x) = \exp[K\{1 - \sin(\omega_0 x + \theta)\}].$$

Thus in step 3 of Algorithm 1, U_i is compared to $\exp[K\{1 - \sin(\omega_0 X_i^* + \theta)\}]$. Equivalently, if unit exponential variates E_i are available, it is faster to compare E_i to $K\{1 - \sin(\omega_0 X_i^* + \theta)\}$, accepting X_i^* if $E_i > K\{1 - \sin(\omega_0 X_i^* + \theta)\}$.

The main computational expense here is generation of the E_i and computation of the sine function, both n^* times. The expense involved in computation of the sine function can be reduced by noting that the point X_i^* is always accepted if E_i is greater than $2K$. This will be a great saving if the cyclic effect is minor (K small). The number of E_i generated can be reduced by noting that if, in one step of the algorithm, E_i is observed to be greater than δ, then $E_i^* = E_i - \delta$ can be used as an (independent) unit exponential variate in the next step. The above procedure can be extended to the case of a trend with two fixed-period cycles, e.g., a time-of-day and a time-of-week effect.

4. SIMULATION OF TWO-DIMENSIONAL HOMOGNEOUS POISSON PROCESSES

The two-dimensional homogeneous Poisson process (of rate $\lambda > 0$) is defined by the properties that the numbers of points in any finite set of nonoverlapping regions having areas in the usual geometric sense are mutually independent, and that the number of points in any region of area A has a Poisson distribution with mean λA; see, e.g. [11], pp. 31-32. Note that the number of points in a region R depends on its area, but not on its shape or location. The homogeneous Poisson process arises as a limiting two-dimensional point process with respect to a number of limiting operations; cf., [8] and [9]. Properties of the process are given by [16]. Applications of the two-dimensional homogeneous Poisson process to problems in ecology and forestry have been discussed by Thompson [21] and Holgate [10]. The model also arises in connection with naval search and detection problems.

In considering the two-dimensional homogeneous Poisson process, projection properties of the process depend quite critically on the geometry of the regions considered. These projection properties are simple for rectangular and circular regions, and make simulation of the homogeneous process quite easy. We consider these two cases separately.

(i) Homogeneous Poisson processes in a rectangle

The following two theorems form the basis for simulation of the two-dimensional homogeneous Poisson process in a rectangle.

THEOREM 2: Consider a two-dimensional homogeneous Poisson process of rate λ, so that the number of points in a fixed rectangle $R = \{(x,y): 0 < x \leqslant x_0,\ 0 < y \leqslant y_0\}$ has a Poisson distribution with parameter $\lambda x_0 y_0$. If $(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$ denote the position of the points of the process in R, labelled so that $X_1 < X_2 < \ldots < X_N$, then $X_1, X_2, \ldots, X_N$ form a one-dimensional homogeneous Poisson process on $0 < x \leqslant x_0$ of rate λy_0. If the points are relabelled $(X_1', Y_1'), (X_2', Y_2'), \ldots, (X_N', Y_N')$ so that $Y_1' < Y_2' < \ldots < Y_N'$, then $Y_1', Y_2', \ldots, Y_N'$ form a one-dimensional homogeneous Poisson process on $0 < y \leqslant y_0$ of rate λx_0.

PROOF: The number of points in an interval on the x-axis, say $(a,b]$ is the number of points in the rectangle bounded by the lines $x = a$, $x = b$, $y = 0$, and $y = y_0$. This number is therefore independent of the number of points in any similar nonoverlapping rectangle bounded on the x-axis by $x = a'$, $x = b'$, i.e., the number of points in the interval $(a',b']$. This establishes the independent increment property for a one-dimensional Poisson process. The Poisson distribution of the number of points in $(a,b]$ follows from the fact that it is equal to the number of points in the rectangle bounded by $x = a$, $x = b$, $y = 0$, and $y = y_0$, and the latter has a Poisson distribution wih parameter $\lambda y_0(b-a)$. An analogous argument shows that the process formed on the y-axis by $Y_1', Y_2', \ldots, Y_N'$ is Poisson.

Conditional properties of the Poisson process in a rectangle are established next. The important thing to note is that while the processes obtained by projection of the points onto the x and y axes are not independent, there is a type of conditional independence which makes it easy to simulate the two-dimensional process.

THEOREM 3: Assume that a two-dimensional homogeneous Poisson process of rate λ is observed in a fixed rectangle $R = \{(x,y): 0 < x \leqslant x_0,\ 0 < y \leqslant y_0\}$, so that the number of points in R, $N(R)$, has a Poisson distribution with parameter $\lambda x_0 y_0$. If $N(R) = n > 0$ and if $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ denote the points, labelled so that $X_1 < X_2 < \ldots < X_n$, then conditional on having observed n points in R, the $X_1, X_2, \ldots, X_n$ are uniform order statistics on $0 < x \leqslant x_0$, and $Y_1, Y_2, \ldots, Y_n$ are independent and uniformly distributed on $0 < y \leqslant y_0$, independent of the X_i.

PROOF: If there are N points in the rectangle, form N vertical strips, from 0 to y_0 and from X_i to $X_i + dx_i$, such that each strip contains only one of the N points. The position of Y_i on the vertical line through X_i is that of an event in a Poisson process of rate λdx_i, given that only one event occurs. But this means that Y_i is uniformly distributed between 0 and y_0. Moreover, this is true irrespective of where X_i occurs; therefore Y_i is independent of X_i. Also, occurrences in all N strips are independent, and therefore Y_i is independent of the other Y_j and X_j positions $j \neq i$. Thus, the Y_i are a random sample of size N from a uniform $(0,y_0)$ distribution, independent of the X_i. Now condition on $N = n$ (> 0); since by Theorem 2 the X_i form a Poisson process they are, by well-known results, order statistics from a uniform $(0,x_0)$ sample and are independent of the fixed size Y_i population; thus the pairs (X_i, Y_i) are mutually independent.

COROLLARY: Denote the Poisson points by $(X_1, Y_1), (X_2, Y_2), \ldots$, where the index does not necessarily denote an ordering on either axis. Conditionally, the pairs $(X_1, Y_1), \ldots, (X_N, Y_N)$ are independent random variables. Furthermore, for each pair, X_i is distributed uniformly between 0 and x_0, independently of Y_i, which is uniformly distributed between 0 and y_0.

From the two theorems, the following simulation procedure is obtained.

ALGORITHM 2: Two-dimensional homogeneous Poisson process in a rectangle.

1. Generate points in the one-dimensional homogeneous Poisson process of rate λy_0 on $(0, x_0]$. If the number of points generated, n, is such that $n = 0$, exit; there are no points in the rectangle.

2. Denote the points generated by $X_1 < X_2 < \ldots < X_n$.

3. Generate $Y_1, Y_2, \ldots, Y_n$ as independent, uniformly distributed random numbers on $(0,y_0]$

4. Return $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ as the coordinates of the two-dimensional homogeneous Poisson process in the rectangle, and n.

Note that generation of the points $X_1, X_2, \ldots, X_N$ in Steps 1 and 2 can be accomplished by cumulating exponential (λy_0) random numbers. Alternatively, after generating a Poisson random number $N=n$ (with parameter $\lambda x_0 y_0$), n independent, uniformly distributed random numbers on $(0,x_0]$ can be ordered; see [13], p. 502.

Another algorithm for generation of the two-dimensional Poisson process in a rectangle can be based on the Corollary to Theorem 3.

(ii) Homogeneous Poisson processes in a circle

The following two theorems form the basis for simulation of the two-dimensional homogeneous Poisson process in a fixed circle of radius r_0.

Fix the origin and initial line of polar coordinates r and θ so that the origin is the center of the circle and the initial line is horizontal. We consider the projection of the points (R_i, θ_i), of the Poisson process circularly onto the r-axis (R_i) and radially onto the circumferential θ-axis (θ_i). The number of points projected onto the r-axis in the interval $(0,r]$, where $r \leqslant r_0$, is the number of points in the circle of radius r and area πr^2; thus the number of points in $(0,r]$ has a Poisson distribution with parameter $\lambda\pi r^2$. Consequently, if the projection process on the r-axis is a Poisson process, it must have integrated rate function $\Lambda(r) = \lambda\pi r^2$, with $\Lambda(0) = 0$.

Similarly, the number of points on the circumferential arc of the fixed circle (radius r_0) from 0 to θ is the number of points in the sector of the circle defined by radial lines at angles 0 and θ; thus the number of points on the arc from 0 to θ has a Poisson distribution with parameter $\lambda\pi r_0^2 \times \dfrac{\theta}{2\pi} = \theta\lambda r_0^2/2$. Accordingly, if the projection process on the θ-axis is a Poisson process, it must have integrated rate function $\Lambda(\theta) = \theta\lambda r_0^2/2$, with $\Lambda(0) = 0$.

We now assert that the projection processes are in fact Poisson processes. Since proofs of these theorems are directly analogous to the proofs of Theorems 2 and 3, they are omitted.

THEOREM 4: Consider a two-dimensional homogeneous Poisson process of rate λ so that the number N of points in a fixed circular area C of radius r_0 and area πr_0^2 has a Poisson distribution with parameter $\lambda\pi r_0^2$. If $(R_1,\theta_1), (R_2,\theta_2), \ldots, (R_N,\theta_N)$ denote the points of the process in C, labelled so that $R_1 < R_2 < \ldots < R_N$, then $R_1, R_2, \ldots, R_N$ form a one-dimensional nonhomogeneous Poisson process on $0 \leqslant r \leqslant r_0$ with rate function $\lambda(r) = 2\pi\lambda r$. If the points are relabelled $(R_1',\theta_1'), (R_2',\theta_2'), \ldots, (R_N',\theta_N')$ so that $\theta_1' < \theta_2' < \ldots < \theta_N'$, then $\theta_1', \theta_2', \ldots, \theta_N'$ form a one-dimensional homogeneous Poisson process on $0 < \theta \leqslant 2\pi$ of rate $\lambda r_0^2/2$.

THEOREM 5: Assume that a two-dimensional Poisson process of rate λ is observed in a fixed circular area C of radius r_0 so that the number of points in C, $N(C)$, has a Poisson distribution with parameter $\lambda\pi r_0^2$. If $N(C) = n > 0$ and if $(R_1,\theta_1), (R_2,\theta_2), \ldots, (R_n,\theta_n)$ with $R_1 < R_2 < \ldots < R_n$ denote the points, then conditional on having observed n points in C, the $R_1, R_2, \ldots, R_n$ are order statistics from the density $f(r) = 2r/r_0^2$ concentrated on $0 \leqslant r \leqslant r_0$, and $\theta_1, \theta_2, \ldots, \theta_n$ are independent and uniformly distributed on $0 < \theta \leqslant 2\pi$, independent of the R_i. These theorems lead to the following simulation procedure.

ALGORITHM 3: Two-dimensonal homogeneous Poisson process in a circular area.

1. Generate n as a Poisson random number with parameter $\lambda\pi r_0^2$. If $n = 0$, exit; there are no points in C.

2. Generate n independent random numbers having density function $f(r) = 2r/r_0^2$ and order to obtain $R_1 < R_2 < \ldots < R_n$.

3. Generate $\theta_1, \theta_2, \ldots, \theta_n$ independent, uniformly distributed random numbers on $(0,2\pi]$.

4. Return $(R_1,\theta_1), (R_2,\theta_2), \ldots, (R_n,\theta_n)$, and n.

Note that the wedge-shaped density $2r/r_0^2$ can be generated by scaling the maximum of two independent uniform (0,1) random numbers.

Direct generation of homogeneous Poisson points in non-circular or non-rectangular regions is difficult. The processes obtained by projection of the points on the two axes are nonhomogeneous Poisson processes with complex rate functions determined by the geometry of the region. However, the conditional independence which is found in circular and rectangular regions (Theorems 3 and 5) for the processes on the two axes is not present. In particular, given that there are n points $(X_1, Y_1), \ldots, (X_n, Y_n)$ in a non-rectangular region, the pairs (X_i, Y_i) are mutually independent, but X_i is in general not independent of Y_i, $i = 1, \ldots, n$. Therefore, it is simpler to enclose the region in either a circle or a rectangle, generate a homogeneous Poisson process in the enlarged area, and subsequently exclude points outside of the given region.

5. SIMULATION OF TWO-DIMENSIONAL NONHOMOGENEOUS POISSON PROCESSES

The two-dimensional nonhomogeneous Poisson process $\{N(x,y): x \geqslant 0, y \geqslant 0\}$ is specified by a positive rate function $\lambda(x,y)$ which, for simplicity, is assumed here to be continuous. Then the process has the characteristic properties that the numbers of points in any finite set of nonoverlapping regions having areas in the usual geometric sense are mutually independent, and that the number of points in any such region R has a Poisson distribution with mean $\Lambda(R)$; here $\Lambda(R)$ denotes the integral of $\lambda(x,y)$ over R, i.e., over the entire area of R.

Applications of the two-dimensional nonhomogeneous Poisson process include problems in forestry and naval search and detection. The use of the process as a model for the pattern of access to the storage subsystem of a computer system will be reported elsewhere. Detection and statistical analysis of trends in the two-dimensional nonhomogeneous Poisson process is discussed by Rantschler [18].

Theorem 1 dealing with thinning of one-dimensional nonhomogeneous Poisson processes generalizes to two-dimensional nonhomogeneous Poisson processes. Thus, suppose that $\lambda(x,y) \leqslant \lambda^*(x,y)$ in a fixed rectangular region of the plane. If a nonhomogeneous Poisson process with rate function $\lambda^*(x,y)$ is thinned according to $\lambda(x,y)/\lambda^*(x,y)$ (i.e., each point (X_i, Y_i) is deleted independently if a uniform (0,1) random number U_i is greater than $\lambda(X_i, Y_i)/\lambda^*(X_i, Y_i)$), the result is a nonhomogeneous Poisson process with rate function $\lambda(x,y)$. The proof is a direct analogue of the proof for the one-dimensional case.

The nonhomogeneous Poisson process with rate function $\lambda(x,y)$ in an arbitrary but fixed region R can be generated by enclosing the region R either in a rectangle or a circle, and applying Algorithm 2 or Algorithm 3. The following procedure assumes that the region R has been enclosed in a rectangle R^*, and that $\lambda^* = \max\{\lambda(x,y): x,y \in R\}$ has been determined; here the bounding process is homogeneous with rate λ^* in the rectangle R^*.

ALGORITHM 4: Two-dimensional nonhomogeneous Poisson process.

1. Using Algorithm 2, generate points in the homogeneous Poisson process of rate λ^* in the rectangle R^*. If the number of points, n^*, is such that $n^* = 0$, exit; there are no points in the nonhomogeneous Poisson process.

2. From the n^* points generated in 1, delete the points that are not in R, and denote the remaining points by $(X_1^*, Y_1^*), (X_2^*, Y_2^*), \ldots, (X_m^*, Y_m^*)$ with $X_1^* < X_2^* < \ldots < X_m^*$. Set $i = 1$ and $k = 0$.

3. Generate U_i uniformly distributed between 0 and 1. If $U_i \leqslant \lambda(X_i^*, Y_i^*)/\lambda^*$, set $k = k+1$, $X_k = X_i^*$ and $Y_k = Y_i^*$.

4. Set i equal to $i+1$. If $i \leqslant m^*$, go to 3.

5. Return (X_1, Y_1), (X_2, Y_2), $\ldots$, (X_n, Y_n), where $n = k$, and n.

It is not necessary that the bounding process have a constant rate λ^*. Theorems 2 and 4 can be generalized to certain cases where the process is nonhomogeneous (cf., [3]), for instance $\lambda(x,y) = \rho(x)\ \psi(y)$. Thus, a tighter bounding process which is nonhomogeneous may possibly be obtained. It is not simple to see how much efficiency could be gained by doing this, as opposed to using a two-dimensional homogeneous Poisson process for the bounding process. Again, as in the one-dimensional case, savings in computing $\lambda(x,y)$ can be obtained by computing its minimum beforehand, and the U_i's can be reused by scaling.

6. COMPARISONS AND CONCLUDING REMARKS

The method of thinning presented in this paper for simulating one-dimensional and two-dimensional nonhomogeneous Poisson processes with given rate function can be carried out in a computationally simple way by using a bounding process which is homogeneous with a rate function equal to the maximum value of the given rate funciton. No numerical integration, ordering, or generation of Poisson variates is required, only the ability to evaluate the given rate function. The thinning algorithm appears to be particularly attractive in the two-dimensional case where there seem to be no competing algorithms.

The thinning algorithm can also be implemented more efficiently at the cost of programming complexity and by using a nonhomogeneous bounding process. In particular the method can be used in conjunction with the special algorithms given by [13] and [15].

It is also possible to extend the method of thinning to simulation of doubly stochastic or conditioned Poisson processes. This will be discussed elsewhere.

REFERENCES

[1] Ahrens, J. and U. Dieter, "Computer Methods for Sampling from Gamma, Beta, Poisson and Binomial Distributions," Computing 12, 223-46 (1974).

[2] Ahrens, J. and U. Dieter, *Non Uniform Random Numbers,* (Technische Hochschule in Graz, Graz, Austria, 1973).

[3] Bartlett, M. S. "The Spectral Analysis of Two-dimensional Point Processes," Biometrika 51, 299-311 (1964).

[4] Cinlar, E., *Introduction to Stochastic Processes,* (Prentice Hall, Englewood Cliffs, New Jersey, 1975).

[5] Cox, D. R., "The Statistical Analysis of Dependencies in Point Processes," in *Stochastic Point Processes,* Ed. P. A. W. Lewis, pp. 55-56 (Wiley, New York, 1972).

[6] Cox, D. R. and P. A. W. Lewis, *The Statistical Analysis of Series of Events,* Methuen, London (1966).

[7] Gilchrist, R., "Some Problems Encountered in Generating Realisations from a Useful Non-Homogeneous Poisson Process," in *Proceedings of the European Congress of Statisticians,* Ed. J. R. Barra et al. (Amsterdam, North Holland, 1976).

[8] Goldman, J. R., "Stochastic Point Processes: Limit Theorems," Annals of Mathematical Statistics 38, 771-79 (1967).

[9] Goldman, J. R., "Infinitely Divisible Point Processes in R^n," Journal of Mathematical Analysis and Applications *17*, 133-46 (1967).

[10] Holgate, P., "The Use of Distance Methods for the Analysis of Spatial Distibutions of Points," in *Stochastic Point Processes*, Ed. P. A. W. Lewis, pp. 122-25 (Wiley, New York, 1972).

[11] Karlin, S. and H. M. Taylor, *A First Course in Stochastic Processes* (Academic Press, New York, 1975).

[12] Lewis, P. A. W., "Recent Advances in the Statistical Analysis of Univariate Point Processes," in *Stochastic Point Processes*, Ed. P. A. W. Lewis, pp. 1-54 (Wiley, New York, 1972).

[13] Lewis, P. A. W. and G. S. Shedler, "Simulation of Non-Homogeenous Processes with Log-Linear Rate Function." Biometrika 63, 501-05 (1976).

[14] Lewis, P. A. W. and G. S. Shedler, "Statistical Analysis of Non-Stationary Series of Events in a Data Base System," IBM Journal of Research and Development 20, 465-82 (1976).

[15] Lewis, P. A. W. and G. S. Shedler, "Simulation of Nonhomogeneous Poisson Processes with Degree-Two Exponential Polynomial Rate Function," IBM Research Report RJ 1953, San Jose, California (1977). To appear in Operations Research.

[16] Miles, R. E., "On the Homogeneous Planar Poisson Point Process," Mathematical Biosciences 6, 85-127 (1970).

[17] Patrow, M. L., "A Comparison of Two Algorithms for the Simulation of Non-Homogeneous Poisson Processes with Degree-Two Exponential Polynomial Intensity Function," M. S. Thesis, Dept. of Operations Research, Naval Postgraduate School, Monterey, California (National Technical Informaiton Service, Springfield, Virgina AD-A047164) (1977).

[18] Rantschler, R. D., "The Detection and Estimation of Trends in Two-Dimensional Poisson Processes," M.S. Thesis, Dept. of Operations Research, Naval Postgraduate School, Monterey, California (National Technical Information Service, Springfield, Virginia; AD-A76136) (1973).

[19] Reyment, R. A., "Geomagnetic Reversal Data Statistically Appraised," Tectonophysics 31, 73-91 (1976).

[20] Steinijans, V., "A Stochastic Point Process Model of the Occurrence of Major Freezes in Lake Constance," Applied Statistics 25, 58-61 (1976).

[21] Thompson, H. R., "Spatial Point Processes with Applications to Ecology," Biometrika 42, 102-15 (1955).

[22] Vere-Jones, D., "Stochastic Models for Earthquake Occurrence, Journal of Royal Statistical Society, B 32, 1-62 (1970).